Name (print): ________________________________

- **INSTRUCTIONS:**
  - Keep your eyes on your own paper, and do your best to prevent anyone else from seeing your work.
  - Do NOT communicate with anyone other than the professor/proctor for ANY reason in ANY language in ANY manner.
  - This is exam is closed book, closed notes, no calculator, and no computer.
  - Turn all mobile devices off and put them away now. You cannot have them on your desk.
  - Write neatly and clearly. What I cannot read, I will assume to be incorrect.
  - Stop writing immediately when told to do so at the end of the exam. I will take 5 points off your exam if I have to tell you multiple times to do so.
  - Academic misconduct will not be tolerated and will be referred immediately to the Emory Honor Council. Penalties for misconduct will be a zero on this exam, an F grade in the course, and/or other disciplinary action that may be applied by the Emory Honor Council.

- **TIME:** This exam has 9 questions on 10 pages including the title page. Please check to make sure all pages are included. You will have 75 minutes to complete this exam.

---

*I commit to uphold the ideals of honor and integrity by refusing to betray the trust bestowed upon me as a member of the Emory community. I have also read and understand the requirements and policies outlined above.*

Signature: ________________________________

---

<table>
<thead>
<tr>
<th>Page:</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points:</td>
<td>12</td>
<td>8</td>
<td>9</td>
<td>12</td>
<td>13</td>
<td>7</td>
<td>9</td>
<td>5</td>
<td>75</td>
</tr>
<tr>
<td>Score:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1. For each of the following terms write a 1-2 sentence explanation/definition.

(a) (2 points) packet

**Solution:** small “sub-unit” of data. information on internet is broken up into and transmitted in packets

(b) (2 points) spear phishing

**Solution:** targeted phishing attacks, usually targeting just a few people or small organization

(c) (2 points) client

**Solution:** computer generating requests for information (usually web pages, email, files, etc) to a server.

(d) (2 points) malware

**Solution:** software which has a malicious intent

(e) (2 points) URL or uniform resource locator

**Solution:** “address” that we’re used to typing into our browser to reach a webpage.

(f) (2 points) zombie computer

**Solution:** computer which has been taken over and can be controlled remotely for malicious purposes.
2. (8 points) In class and in your readings, the data-information-knowledge continuum was discussed. For each of the following scenarios, give an example of each part of this continuum and illustrate how data, information, and knowledge are inter-related.

(a) A biological conservation organization which is studying endangered plant species.

**Solution:**

- **Data:** latitude/longitude coordinates of endangered species for a specific sampling area. Perhaps time/date data as well for when plant was logged.
- **Information:** data plotted on terrain map for specific period of time
- **Knowledge:** being able to see that plant species is increasing/decreasing over time in a particular space

(b) A utility company which wants to bring down average customer power usage during the summer.

**Solution:**

- **Data:** kW usage for a specific address, timestamped to allow indicate time/date information
- **Information:** calculation of customer’s average kW usage per day
- **Knowledge:** comparison’s between customer’s usage and neighbor’s usage or between customer’s usage and other’s who live in same size house/apartment.

Most common errors: confusing data and information. Many people listed information in place of data without clearly specifying where data would come from.
3. Given the spreadsheet below, evaluate each of the following formulas. If the formula will generate an error, you can simply write “error”. You do not need to specify the type of error. **Use quotation marks to clearly differentiate text from numbers.**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>05/20/99</td>
<td>03:35:00 AM</td>
<td>TEXT HERE</td>
</tr>
<tr>
<td>2</td>
<td>03/03/12</td>
<td>04:47:00 PM</td>
<td>cot rot hot</td>
</tr>
<tr>
<td>3</td>
<td>07/18/05</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) (1 point) =MONTH(A2)  
(b) (1 point) =DAY(A3)  
(c) (1 point) =DATE(YEAR(TODAY()), MONTH(A3), DAY(A1))  
(d) (1 point) =IF(MONTH(TODAY()) > MONTH(A1), "Yes", "No")  
(e) (1 point) =TIME(14,12,21)  
(f) (1 point) =HOUR(B2) - MIN(B1)  
(g) (1 point) =SUBSTITUTE(UPPER(C2),"O", "a")  
(h) (1 point) =LEN(C1)  
(i) (1 point) =SEARCH("e", C1)
4. (12 points) Each of the following items is an issue that a business must address in order to design, develop, and deploy its data systems. For each item listed below, give an example which clearly demonstrates why the issue is a crucial one for businesses to consider in their data management strategy.

(a) Data sourcing

**Solution:** See reading 11.5 for more information.

Where does data come from? Example: sales data for reports or analysis can be gathered from point of sale or transaction systems.

(b) Data quality

**Solution:** Is data accurate and complete? Example: if the data in our system comes from handwritten surveys which must be manually entered, we need to have some sort of check to make sure the person entering the data isn’t making mistakes in transcription.

(c) Data hosting

**Solution:** Where data is housed. Example: Google chooses to host its servers in a geographical area where there is 1) cheap, plentiful power and 2) fast backbones for connectivity.
5. (4 points) List two arguments for and two arguments against net neutrality.

**Solution:** For:
+ Net neutrality encourages innovation since smaller companies w/ innovative ideas won’t have to pay money to compete against larger companies.
+ Limiting internet speeds could skirt dangerously close to censorship or freedom of speech issues.
+ Net neutrality is the current state of things, and telecoms should not be allowed to change this.

Against:
- Since telecoms spend money on infrastructure, they should have the right to dictate the use of this infrastructure to profit.
- Telecoms will have little incentive to invest in better/faster infrastructure if they can’t profit more from it.
- Capitalism/free market economic principles would dictate this approach.

See reading 12.3 for discussion.

Common mistakes: reversing the pro/con sides of things. The idea that net neutrality is about “blocking” content (legal or illegal). Net neutrality is about establishing (or not) a two tier system: one faster and one slower, but neither of which completely blocks content.

6. (3 points) When visiting websites, we can type in a human-readable address like http://www.google.com or we can type in an address that looks something like http://74.125.224.72. Explain why either of these addresses will work. What system or service allows us to translate human readable addresses to numbers like those above?

**Solution:** Human readable addresses (URLs) (eg google.com) map to a specific IP addresses (the numbers). All information is delivered using the IP address, but DNS (Domain Name Service/Servers) allows us to enter URLs which are meaningful/recognizable to humans. DNS then translates the text to the numbers for us.

7. (6 points) Consider the following web address:
   http://news.emory.edu/stories/2013/03/aids_vaccine_200_bicycle_ride/campus.html
Split this address into its components listed below.

- file: campus.html
• top level domain: edu
• host name: news
• path: stories/2013/03/aids...ride
• protocol: http
• domain: emory
8. (7 points) Define the term “social engineering” and give an three examples of how it could be used to compromise an organizations information security.

Def’n: Using a con of some sort which exploits human nature to compromise security in some manner.

Many possible examples, but some are:
1) Holding heavy boxes and running for the door in order to convince someone to hold open a door that usually requires a badge to open and “tailgating” through the door.
2) Acting with an air of authority and pretending to belong in a secure area.
3) Pretending to be a new employee whose badge or log-in credentials aren’t working for some reason.
4) spear phishing can be considered a type of social engineering.

See reading 13.3 for other examples.

Most common errors: listing 3 examples which were all small variants of one type of social engineering. Example: impersonation 1) Impersonating someone’s boss..., 2) impersonating a tax auditor ...3) impersonating a security guard...

Other common mistakes: treating social engineering as simply a way in which computers could be compromised. Needed to have some aspect of the “con” for full credit. Defining social engineering as something that could only happen online or over the internet.
9. Consider the database illustrated on the next page which keeps track of employees and what projects they are working on. In this picture, a relationship of “Many” has been illustrated with an infinity symbol.

(a) (2 points) For the Projects table, list one field and one record.

**Solution:** field is a piece of stored data: Proj_ID, date assigned, date completed, or Description were acceptable answers.

record is a complete set of fields: listing any row from table is acceptable.

(b) (3 points) You need to be able to track which employees are project managers. How would you modify the database to include this information? Draw your answer on the next page (page 9).

**Solution:** Easiest way to do this is to add a field to the EMPLOYEES table which stores whether or not the employee is a manager.

Common mistake: adding a “Project Manager ID” field to the EMPLOYEES table instead. This introduces another primary key into the table (and you don’t really need a unique ID in this case; you already have one in this table: EmployeeID).

(c) (4 points) Assume a project can have 1 and only 1 manager assigned to oversee it. What is the best way to incorporate this requirement into the database? Draw your answer on the next page (page 9) and justify why your method is the best way to accomplish this requirement.

**Solution:** Easiest way to do this is to include a field in the PROJECTS table for a Project Manager Employee ID. This ID would be linked to the EMPLOYEES table as a foreign key. For this part, it would have a 1 to many relationship. That is, a project can have only 1 manager, but 1 employee can manage many projects.

Several people built an extra table similar to the ASSIGNMENTS table, but for project managers. Since we’re guaranteed that a project will have only 1 manager, the most efficient solution would be to just store the p.m employee id as part of each record in the projects table.

Some people just stored an EmployeeID in the PROJECTS table to answer both parts b and c. This isn’t really ideal since whether or not an employee is a manager (or has the credentials to manage) is part of the employee data rather than the project data.
Common mistake: setting up relationships and a structure which meant that a manager could only manage one project. The problem states that a project can have only 1 manager, but nothing about how many projects a manager can supervise.

(d) (5 points) Now assume that a project can be managed by more than 1 manager. In other words, a project may have multiple managers and multiple workers assigned to it. Modify the database to accommodate this managerial structure. Draw your answer on page 10 and justify why your way is the best way to modify the database. If you add new table(s), you do not need to populate it/them with sample data. However, be sure to indicate relationships to other tables correctly.

**Solution:** There are two ways to do this. The first is the simplest: simply add a MANAGER field to the assignments table. This field classifies each employee’s role on a project as either a manager or not. This is a very flexible solution since employee may manage some projects and work on others without managing them. Additionally, many employees can be assigned to a project in the role of manager.

You could also build a separate table, similar to the ASSIGNMENTS table which would map employee id’s (in this case, of people who were managers) to projects. This introduces some extra information shuffling since employees could appear in both the ASSIGNMENTS and the new table. They could either be “just an employee” on a project and thus be a record in the ASSIGNMENTS table and perhaps be a manager on another project(s). However, this would implement a correct solution.